

Forensic Analysis and Genetic Source Partitioning Model for Portland Harbor Contaminated Sediments

T. A. T. Aboul-Kassim* and K. J. Williamson

Environmental Analysis and Impact Assessment Group, Department of Civil, Construction and Environmental Engineering,
College of Engineering, Oregon State University, 202 Apperson Hall, Corvallis, OR 97331, USA

ABSTRACT. Point and non-point source pollution to aquatic systems pose many challenges in maintaining ecosystem integrity. Aquatic sediments in many rivers, lakes, and harbors have been directly or indirectly contaminated by various types of chemical compounds. The Portland Harbor, a representative example of a contaminated sediment study, has played a key role in economic development of the City of Portland and the State of Oregon (USA) for decades. In addition, it has been a valuable resource for recreation, fishing, and navigation. Because of its current contamination, the United States Environmental Protection Agency (EPA) has placed the harbor onto the National Priorities List by its designation as a federal superfund site. Therefore, this paper aims at demonstrating a unique forensic analysis and genetic source partitioning model that would help characterize the extent of organic contamination and build up an ultimate sampling scheme to be adopted by the Oregon Department of Environmental Quality (DEQ) and EPA to aid in the final phase of feasibility study and remedial investigation of the harbor. Several sediment samples, collected from the Portland Harbor, are analyzed, characterized, and fingerprinted in terms of their hydrocarbon (HC) organic molecular marker (MM) signatures in this study. The distributions, chemical structures, and applicability of such MMs in determining characteristic group(s) representative for the study area are discussed and evaluated in this paper using the HC multi-tracer environmental forensic MM approach. Homologous long chain *n*-alkanes (C₁₆-C₃₈), carbon preference index (CPI), unresolved complex mixtures (UCM), and MMs such as pristane, phytane, tricyclic (C₁₉-C₂₉) and tetracyclic (C₂₄, C₂₈ and C₂₉) terpanes, 17 α (H), 21 β (H)-hopanes (C₂₇-C₃₅), 5 α (H),14 β (H),17 β (H)-steranes (C₂₇-C₂₉) with a minor amounts of 5 α (H),14 α (H),17 α (H)-steranes and 13 α (H),17 β (H)-diasteranes are found to be the most suitable indicators to differentiate between petroleum hydrocarbon- from non-petroleum hydrocarbon-containing sediment particles. In contrast, high temperature thermogenic/pyrolytic-derived compounds are indicated by a specific group of polycyclic aromatic hydrocarbons (PAHs). This group, ranging from phenanthrene to dibenzo(a,e)pyrene with different alkyl-substituted PAH series, is considered to be combustion products from fossil fuel. Extended Q-mode factor analysis and linear programming technique are also performed to: (a) examine the variations in the hydrocarbon MM data set, (b) group the data into specific associations (i.e., end-members), and (c) find statistically significant clusters in the data set to help assess and identify the various hydrocarbon pollution sources and original compositions reflecting aquatic sediment impact to the Portland Harbor.

Keywords: sediment, forensic analysis, genetic source partitioning model, chemometrics

1. Introduction

Implementing the environmental analysis and impact assessment (EAIA) concept to track down contaminated sediments is intimately related to applying an integrated process for identifying and evaluating the fate and transport consequences of pollutants of concern in the aquatic environment. In many rivers, lakes, or harbor areas, bottom sediments are mostly composed of complex organic mixtures of various contaminants, which are of environmental concern due to their high concentrations, toxicity, persistence, and bioaccumulation rate (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002). Contaminated sediments can pose big potential hazard to large groups of aquatic species. The situation could be even worse especially if the detected

pollutants are toxic, carcinogenic, mutagenic and/or endocrine disruptive and the impact zone is close to the human beings' communities.

In order to distinguish the different anthropogenic sources of a contaminated sediment environment, it is important that the sediment composition from various sources be identified, the proportion of sediments from each source be determined, and the degree of contamination of each sediment type be calculated. The most straightforward approach to performing such an evaluation would be to isolate sediments from individual sources (i.e., end-members), to determine their composition analytically, and to assess their relative importance in the sediment. Most bottom sediments, however, are intimate mixtures of several sediment types. The fine grain size of the sediments generally precludes direct determination of the actual composition of sediments from various sources or the relative importance of the various sources in sedi-

*Corresponding author: tarek.kassim@oregonstate.edu

ments. For these reasons, many researchers have attempted to partition bulk chemical analyses of bottom sediments with respect to their source terms. The solution to fulfilling such a partitioning task requires: (a) an estimate of the number of important end-member sources represented in the data set being studied, (b) the determination of the chemical compositions of the end-member sediments, and (c) the determination of the absolute abundance of each sediment source in each sample.

Two approaches were used to determine the chemical composition of proposed end-members: direct chemical analysis of sediment end-members and modeling based on such chemical analysis (Aboul-Kassim and Simoneit, 1995, 1996, 2001). Studies using the first approach seek to determine end-member compositions directly by analysis of samples presumed to be pure end-members or of sediment fractions that are physically separated. Others use differential chemical leachates that extract some components of the sediment presumed to represent end-members (Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002). Such analytical approaches are complicated by the intimate mixture of phases and the occasional chemical alteration of phases due to diagenesis or re-crystallization. In addition, analytical separation techniques like chemical leaching may strip trace organics which are adsorbed on fine-grained phases. In the second approach, the analyses were normally combined with interferences obtained from multivariate statistical analysis to further refine the relationships between contaminants in a given source (Aboul-Kassim and Simoneit, 2001). All of these approaches require a series of assumptions about the samples representing the end-members and/or about the chemistry of the sources.

Several studies used multivariate analysis based on bottom sediments data sets to determine quantitatively the contribution of each contaminant from various sediment sources, (Aboul-Kassim and Simoneit, 2001; Dymond, 1981; Full et al., 1981; Grant, 1990; Imbrie and Van Andel, 1971; Klován and Imbrie, 1971; Malinowski, 1991; Miesch, 1976; Rapp, 1991; Tysklind et al., 1992). Within such studies, they generally assumed that sediments are made up of a few end-member components and that as long as the area and time interval being considered are reasonably restricted, the sources are constant in composition. Early partitioning efforts; emphasized the use of inner-contaminant ratios to determine the importance of each end-member in a bulk sediment sample. Another approach to fulfilling partitioning task uses normative analysis and linear programming (Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002; Dymond, 1981; Heath and Dymond, 1981). The linear programming technique is a significant improvement over other normative analysis techniques because there are more equations (source-contaminant relationships) than unknowns (percentage of each source in the sample) and

the system can be solved explicitly.

Accordingly, the main objective of the EAIA approach presented in this paper is to solve the partitioning problem that requires fewer assumptions than before. This unique EAIA partitioning model, which is specific for the Portland Harbor bottom sediments, consists of two main inter-related approaches: (a) Forensic Analysis, and (b) Genetic Source Partitioning Model.

Environmental forensic analysis (EFA) is a scientific methodology developed for analyzing and identifying contamination-related and other potentially hazardous environmental contaminants, and for determining their sources. It combines experimental analytical procedures with scientific principles derived from several science- and engineering-disciplines. It can also provide a valuable tool for obtaining scientifically proven data when applied to investigations of sediment contamination. In general, EFA is effective in many liability cases that aim to successfully identify the party(s) responsible for contamination (i.e., actual sources). In most EAIA cases, the presence of chemicals is determined following standard US EPA methods. For instance, EPA Method 418.1 (for analysis of total petroleum hydrocarbons) or Method 8015 (for determining the presence of petroleum in the diesel or gasoline range) is generally adequate for documenting the occurrence of petroleum contamination at a certain site. However, these methods are unsuitable for identifying the types and sources of petroleum in complex contaminated situations. These standard EPA methods are: (a) not intended for petroleum product identification, (b) not tailored for the analysis of the key diagnostic chemical compounds that comprise petroleum, and (c) do not provide sufficient chemical data to perform defensible data analysis for source identification and product differentiation.

An accurate and defensible forensic analysis approach requires answering the following key questions: (a) What are the product types present due to contamination? (b) What are the potential sources of contamination? and (c) Can these potential sources be linked to their original sources? Answers to these questions require the use of sophisticated contaminant-specific methods of chemical analysis together with advanced data analysis, statistical/mathematical modeling and visualization techniques. Successful fingerprinting involves the design and implementation of an investigation including sampling, analytical and interpretation strategies. The selection of appropriate marker compounds that will differentiate contaminant sources is central to designing an analytical program. Because hydrocarbons, for instance, have a variety of sources, the most important objective in analytical selection is to identify specific markers for both the released/leached hydrocarbon mixture and other potential sources. Markers must have the attributes of uniquely identifying the released/leaching contaminants from other sources and resistance to alteration (i.e., weathering) over time. Most often, the appropriate markers are not known at the start of a fingerprinting study, but are identified

during the characterization processes. Successful contamination fingerprinting usually requires analysis of more than one target analyte group.

Although a major element in an environmental forensics investigation is the chemical fingerprinting that is designed to identify contaminants; however, fingerprinting alone is not sufficient enough to provide answers to questions of source location and legal responsibility. Accordingly, incorporation of Chemometrics (i.e., the numerical analysis of chemical data) to a forensic analysis data set would synthesize all the complex information to make it easily visualized. Such methods were supplanted by a number of powerful data interpretative tools, e.g., graphical techniques that examine relationships among marker compounds and advanced statistical multivariate analysis for examining these relationships (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1995, 1996, 2001; Aboul-Kassim and Williamson, 2002, Grant, 1990; Malinowski, 1991; Rapp, 1991; Tysklind et al., 1992).

This paper, which is the first presentation of a series of companion studies, aims to demonstrate a unique forensic analysis and genetic source partitioning model for analyzing contaminated aquatic sediments in the Portland Harbor, Oregon, USA (see Figure 1). This would help characterize the extent of organic contamination and build

up an ultimate sampling scheme to be adopted by the Oregon Department of Environmental Quality (DEQ) and EPA to aid in the final phase of feasibility study and remedial investigation of the harbor. To fulfill the entire spectrum of the study, it requires: (a) determining the nature and extent of organic contamination in the harbor, (b) characterizing the various contaminants classes, and (c) partitioning the harbor sediments according to their probable input sources. To determine the number of end-members specific for the Portland Harbor bottom sediment, an extended Q-mode factor analysis is used to identify the principal sources of variation (i.e., end-members within the data set). To determine the composition of the end-members, the extended Q-mode factor analysis is combined with a new vector rotation scheme to obtain chemically reasonable compositions for the end-members and the relative importance or concentration of each end-member in each sample. The compositions determined by this new technique can also be used as inputs to supporting partitioning techniques, like linear programming, for differentiating concentration of each end-member within each sediment sample of the harbor. Once the EAIA task can be successfully set up for handling the organic contaminant multi-tracer model in this study, the application of the same model to other groups of contaminated sediment sites may be anticipated

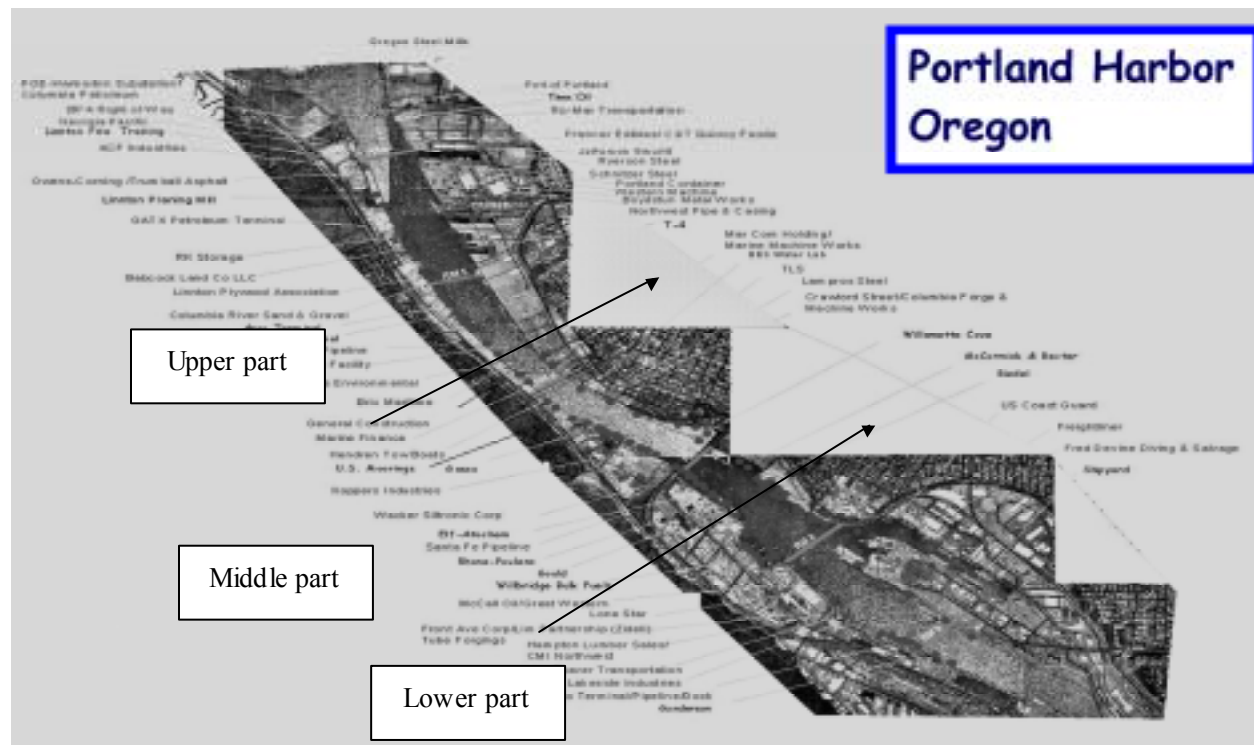


Figure 1. The Portland Harbor, Oregon.

in the future to aid in further comparison, dissemination, and expansion of data bases leading to the provision of a firm basis in support of possible legislation program.

2. Materials and Methods

2.1. Site History and Description

The Willamette River is the 10th largest river in the USA and drains the heart of the State of Oregon. In the 1930s, the Willamette was so polluted that fish were dying and the water was no longer safe for human use. Decades of local effort resulted in significant improvements to water quality and, by the 1970s, the Willamette success became a model for its accomplishment of environmental restoration. Recently, another critical milestone in the history of the Willamette River has been approached. A study performed by the EPA and the Oregon DEQ found that the aquatic environment in the lower river between Sauvie Island and Swan Island (Oregon) are contaminated with various pollutants (Aboul-Kassim and Williamson, 2002). As a result, EPA (in late 2000) has placed this 6-mile stretch of the river, referred to as the Portland Harbor (Figure 1), onto the National Priorities List by its designation as a federal superfund site.

The Portland Harbor has played an important role in economic development of the City of Portland and the State of Oregon for decades. It has been a valuable resource for recreation, fishing, and navigation. Industrial uses of Portland Harbor began in the mid-1800s when the first wharves were constructed to support international and inter-coastal steamship service. In 1868, the first river dredging was conducted and, since then, the Willamette River has been continually dredged for navigational purposes. In 1996, more than 28 million tons of goods were exported through the Columbia and lower Willamette River. Approximately 3.8 million tons of commodities were imported during the same year, including alumina, limestone, cement, and crude salt. Some of the historical and/or current industrial operations along the Portland Harbor include (see Figure 1): marine construction; bulk petroleum product storage and handling; construction material manufacturing; oil fire-fighting training activities; oil gasification plant operations; pesticide/herbicide manufacturing; wood treating operations; agricultural chemical production; battery processing; liquid natural gas plant operations; hazardous waste storage; chlorine production; ship loading and unloading; ship maintenance, repair, and refueling; rail car manufacturing; and metal scrapping and recycling

2.2. Sampling

Bottom sediments were collected by coring from various locations representing the upper (4 samples), middle (2 samples) and lower (2 samples) parts of the Portland Harbor (see Figure 1). The top 10 cm of the cores were homogenized and extracted for environmental analysis. The sampling program covered appropriate regions both in the har-

bor and in the entire length of 6 miles.

2.3. Forensic Analysis

The next few sections explain the comprehensive forensic analysis protocol used for the characterization and identification of the Portland Harbor contaminated sediments.

2.3.1. Extraction and Separation

An extraction protocol developed and revised in this study (Aboul-Kassim and Simoneit, 1995; 2001) was applied for the qualitative and quantitative analyses of different organic classes of compounds found in the bottom sediments of the Portland Harbor. In brief, bottom sediments were extracted in a Soxhlet apparatus with methylene chloride-methanol (2:1). All the extracts were concentrated to 2 ml and hydrolyzed overnight with 35 ml of 6% KOH/methanol. The corresponding neutral and acidic fractions were successively recovered with *n*-hexane (4 × 30 ml), the latter after acidification (pH 2) with 6N HCl. The acidic fractions, previously reduced to 0.5 ml, were esterified overnight with 15 ml of 10% BF₃/methanol. The BF₃/methanol complex was destroyed with 15 ml of water, and the methyl esters were recovered by extraction with 4 × 30 ml of *n*-hexane. The neutrals were fractionated by long column chromatography. The following fractions were collected: (i) 45 ml of *n*-hexane (aliphatic hydrocarbons), (ii) 25 ml of 10% methylene chloride in *n*-hexane (monoaromatic hydrocarbons), (iii) 40 ml of 20% methylene chloride in *n*-hexane (polycyclic aromatic hydrocarbons "PAHs"), (iv) 25 ml of 50% methylene chloride in *n*-hexane (esters and ketones), (v) 25 ml of methylene chloride (ketones and aldehydes), and (vi) 50 ml of 10% methanol in methylene chloride (alcohols). The last fraction and an aliquot of the total extract were derivitized prior to gas chromatographic-mass spectrometric (GC-MS) analysis for further qualitative molecular examination by silylation with bis-(trimethylsilyl) trifluoroacetamide. A recovery experiment for the long column chromatography was carried out using several deuterated standards. Here, we only present the results of fractions 1 to 3 that represent the aliphatic and aromatic hydrocarbons composition of sediment particles.

2.3.2. Instrumental Analyses

High-resolution gas chromatography (GC) was conducted on a Hewlett Packard (HP) 5890A gas chromatograph, equipped with a split/splitless capillary injection system and a flame ionization detector (FID). The samples were analyzed in the splitless mode using a fused silica capillary column (30 m × 0.25 mm i.d, DB-5, 0.25 μm film thickness, J & W Scientific) and helium as carrier gas. The analog signal was monitored and/or integrated with an HP 3393A integrator. The GC conditions were: FID 300 °C, injector 300 °C, oven temperature initially 65 °C, programmed to 200 °C at 4°C/min, isothermal at 290 °C (60 min). The gas chromatography-mass spectrometry (GC-MS)

analyses were performed with an HP GC (identical column with initial temperature 50 °C, isothermal 6 min, programmed at 4°C/min to 310°C, isothermal 60 min) interfaced directly to an HP quadrupole mass spectrometer (electron impact, emission current 0.45 mA, electron energy 70eV, scanned from 50 to 650 daltons).

2.3.3. Identification and Quantification

Compound identification was based on comparison with the GC retention times and/or mass fragmentation patterns of standard reference materials. With the help of the Library, molecular marker identification was tabulated as follows: (a) *Positive*, when the sample mass spectrum, authentic standard compound mass spectrum, and their retention times agreed well; (b) *Probable*, same as above except no standards are available, but the sample mass spectrum agreed very well with the standard library; (c) *Possible*, same as above except that the sample spectrum contained information from other compounds but with minor overlap; and (d) *Tentative*, when spectrum contained additional information from possibly several compounds with overlap. Identification and response factors of pollutants were determined using a suite of standard compounds.

Molecular marker identification was achieved using the following standard mixtures injected in both GC and GC-MS: (a) A series of normal alkanes ranging from *n*-C₁₀ to *n*-C₃₆; (b) Regular isoprenoids as pristane and phytane; and (c) A suite of about 30 polycyclic aromatic hydrocarbons (PAHs), including: naphthalene; methyl- and dimethyl-naphthalene; fluorine; 9-methylfluorene; dibenzo- thiophene; phenanthrene; 3-, 2-, 9- and 1-methyl- phenanthrene; anthracene; fluoranthene; pyrene; 2,3-benzofluorene; 1,1'-binaphthalene; benz(a)anthracene; chrysene; benzo(b+k)fluoranthene; benzo(e)pyrene; benzo(a)pyrene; perylene; 9,10-diphenylanthracene; dibenz(ah)anthracene; benzo(ghi)perylene; Anthanthrene; coronene; and dibenzo(ae)pyrene.

2.3.4. Quantification

Quantification was based on the application of perdeuterated compounds (e.g. *n*-C₃₂D₆₆ and d₁₀-pyrene) as internal standards. In order to correct for detector response, sets of relative response factors were determined for every fraction from multiple injections. Molecular markers were determined and quantified by GC-MS.

2.3.5. Organic Carbon Analysis

Total organic carbon (TOC) analyses were carried out for all sediment samples using a Carlo Erba NA-1500 CNS analyzer. The contaminants concentrations were calculated relative to their TOC contents.

2.4. Genetic Source Partitioning Model

Contaminants data and several chemical indices representing the Portland Harbor bottom sediments were statistically examined using the SAS[®] Statistical Package

Version 6.12 (SAS, 1997). Extended Q-mode factor analysis and linear programming technique were used to construct a source partitioning model specific for the Portland Harbor region (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001). The objectives of the statistical analyses and mathematical modeling were to define the organic molecular marker (MM) characteristics of the Portland Harbor sediments, and assess their sources. The principles of some of these techniques are discussed in the next few sections.

2.4.1. Extended Q-Mode Factor Analysis

Factor analysis has extensively been used in Environmental Chemistry (e.g. Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1995, 2001; Grant, 1990; Malinowski, 1991; Rapp, 1991; Tysklind et al., 1992). Q-mode factor analysis is based on grouping a multivariate data set based on the data structure defined by the similarity between samples. It is devoted exclusively to the interpretation of the inter-object relationships in a data set, rather than to the inter-variable (or covariance) relationships explored with R-mode factor analysis. The objective of Q-mode factor analysis is analogous to geochemical partitioning models that seek to determine the absolute abundance of the dominant components (i.e., MMs) in environmental samples (Aboul-Kassim and Simoneit, 2001, Imbrie and Van Andel, 1971). It provides a description of the multivariate data set in terms of a few end-members (i.e., associations or factors, usually orthogonal) that account for the variance within the data set. A factor score represents the importance of each variable in each end-member. The set of scores for all factors makes up the factor score matrix (Klován and Imbrie, 1971). The importance of each variable in each end-member is represented by a factor score, which is a unit vector in *n* (number of variables) dimensional space, with each element having a value between -1 and 1, and the sum of the squared elements equal to 1.00 (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1995). The relative importance of each end-member factor in each sediment sample is its factor loading value. The complete set of factor loadings describing each sediment sample in terms of its end-members is the factor-loading matrix.

Q-mode factor analysis defines the similarity of objects by considering the component proportions. The method searches elements in the *A* matrix for the most divergent objects, represented by the pure component concentrations or those constituted by a significant proportion of these components, which can be represented as vertices of a concentration simplex. The other data set objects are linear combinations of the divergent ones. The contribution of each object is obtained by an eigen-analysis of a real symmetric matrix obtained from the data matrix. The measure of similarity used is the cosine theta (*cos θ*) matrix, i.e. the matrix whose elements is the cosine of the angles between all sample pairs (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1995, 2001; Imbrie and Purdy, 1962). For two objects, *n* and *m*, *cos θ* is calculated by:

$$\cos \theta_{nm} = \frac{\sum_{j=1}^q a_{nj} a_{mj}}{\sqrt{\sum_{j=1}^q a_{nj}^2 \sum_{j=1}^q a_{mj}^2}} \quad (1)$$

For positive a elements this index varies from zero, no similarity, to one, identity.

The mathematical procedure starts by calculating $\cos \theta$ for all pairs of objects in the data set of matrix A . The first step normalizes the A matrix rows, pre-multiplying A by an $n \times n$ diagonal matrix D^{-1} , the inverse of the D matrix. The principal diagonal of matrix D is composed of square roots of the sum of squares of the row vector elements of A ,

$$W_{(n \times p)} = D^{-1} A \quad (2)$$

The similarity or association matrix is defined by

$$H = WW' = D^{-1} AA' D^{-1} \quad (3)$$

and can be approximately expressed as the product of the score, $T_{(n \times q)}$, and the loading matrix, $P'(p \times q)$, with q being the approximate rank of matrix W . The T matrix, determined by Imbrie and Purdy (1962), does not furnish a set compositionally distinct objects. One way of resolving this problem is by means of varimax and oblique rotations.

Simple Q-mode factor analysis fails to provide a direct solution to the partitioning problem, however, because: (a) the vectors generated by factor analysis are not composition vectors and thus cannot be used to indicate the absolute composition of the end-members, (b) the factor scores only give a relative measure of the importance of each variable in each end-member and also reflect any scaling done on the data set prior to the analysis (such as transforming variable values to percent of range or normalizing variables to equal means), (c) the factor score matrix can contain negative values (for compositions of geochemical end-members this is an unreasonable condition), (d) the factor leading matrix indicates only the relative importance of each end-member and not an absolute abundance, and (e) the factor loading matrix also commonly contains negative values.

An extension of Q-mode factor analysis (Klovan and Miesch, 1976; Miesch, 1976) provides a solution to the first, second and fourth problems listed above, those associated with obtaining absolute compositions of the end-members themselves. Normally when factor analysis is used there is no restriction placed on the sum of the variables within a sample. For example, with geochemical data, if the data are expressed in wt% or in ppm there is no requirement that the sum of the variables in a simple equal 100% or 1,000,000 parts to do the analysis, because the analysis is based on ratios. If the data are closed by

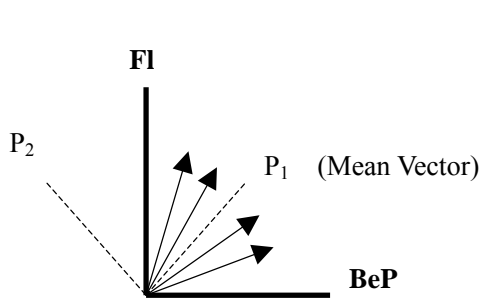
summing the variables in each sample to a constant value, the factor scores can be converted into actual composition of the original variables (Miesch, 1976). Also, with a closed data set, the absolute abundance of each end-member factor in each sample can be calculated.

The problem of negative values in the factor loading matrix can be eliminated if a "varimax" rotation is used in the factor analysis (Imbrie and Van Anandel, 1962). This rotation of the factor vectors ensures that the absolute abundance of each end-member factor in each sample is greater than or equal to zero. The reason for this can be seen by a two-dimensional illustration of the "varimax" rotation (Figure 2b). The first operation in a Q-mode factor analysis points a vector in the direction of maximum variability of the data set (Figure 2a). This direction usually ties near the vector representing the mean value for all variables. A second vector points in the direction of the next greatest variability and is restricted to be orthogonal to the first vector. The addition of principal axes is continued until the desired amount of the variability of the data set is described. As can be seen in the two dimensional example in Figure 2a, the second axis is constrained to point in an extreme direction compared to the sample compositions and is usually a chemically unreasonable composition (in this case an end-member with a negative concentration of BeP, see Figure 2). The "varimax" rotation rotates the principal component axes so that the variability within the data set explained by each axis is maximized (with the restriction that the axes remain orthogonal). This rotation brings the end-member axes closer to real simple compositions (Imbrie and Van Anandel, 1962; Aboul-Kassim and Simoneit, 2002). Thus, varimax rotation rigidly rotates the vectors of the T matrix until they coincide with the most divergent vectors in space. The order of the T matrix for rotation is equal to the number of simplex vertices, which is determined from the number of significant eigen values. The T matrix is then rotated to produce a new matrix, F , where

$$F = TR \quad (4)$$

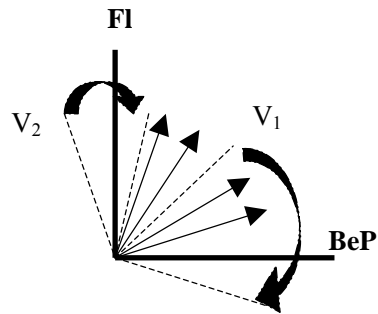
with $R_{(q \times q)}$ being the transformation matrix and $F_{(n \times q)}$ being the varimax weights. Each row of F corresponds to an object or a linear combination of objects and each column represents a factor. Also, as can be seen in Figure 2b, all samples can now be described in terms of positive or zero contributions of the end-member factors.

Thus, the combination of using a data set in which the sum of the variables in each sample is constant and of using a varimax rotation results in a factor analysis which describes each sample in terms of the absolute abundance of end-member factors whose compositions are given in terms of concentrations of the original data variables. Unfortunately, as can be seen from Figure 2b, the compositions of these varimax end-members can and usually do contain large negative values for some variables (which may reflect inverse relationships) and thus cannot represent real geochemical end-member compositions.



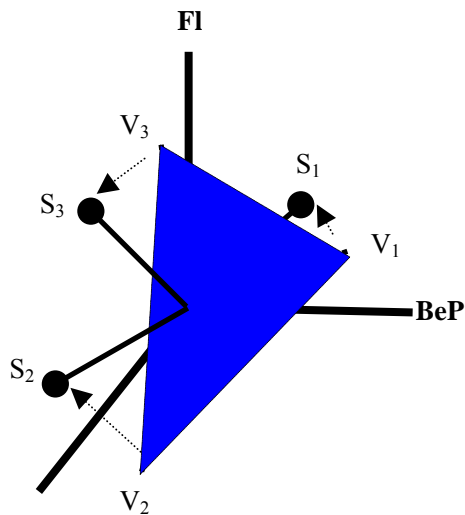
(A) Principal Components

Example of principal components (PCs) determined for a hypothetical two variable system with Fluorene (Fl) and Benzo(e)pyrene (BeP). Solid vectors with arrowhead represent Fl/BeP of data; P_1 is the first PC or mean vector; P_2 is the second PC and is constrained to be orthogonal to P_1 .



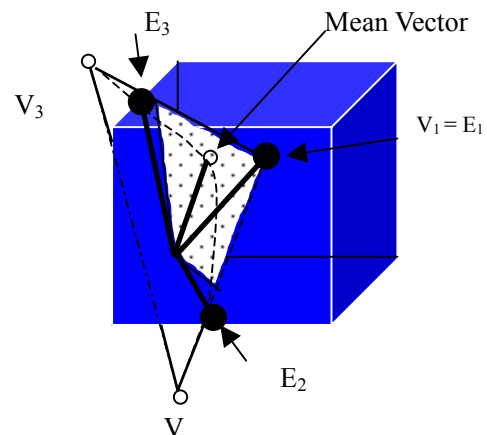
(B) Varimax Rotation

Example of a "varimax" rotation of the system in Fl. The principal components have been rotated so that the sum of squares within the data set explained by each axis is maximized.



(C) Oblique Rotation

Example of an "oblique" vector rotation for a hypothetical three variable system with Fl, BeP and Coronene (Cor). The dotted lines indicate and represent a set of principal component (PC) vectors " V_1 , V_2 , and V_3 ". V_3 has a negative value for BeP, a chemically unreasonable composition. The oblique rotation rotates each of the PC vectors to the sample vector (solid lines with data points) nearest it, resulting in positive values for all variables.



(D) New Rotation

Example of the new rotation scheme as applied to the three variable system shown in Figure 2C. Open dots labeled V_1 - V_3 represent the original principal component (PC) vectors. Light dotted lines with arrows show the path of rotation of the PC axes which are rotated toward the mean until they intersect the positive vector space. Solid lines with black dots labeled E_2 and E_3 show the resulting end-member vectors. V_1 is not rotated since it is already in the positive vector space.

Figure 2. Different vector rotation techniques.

To eliminate the problem of end-members with negative, variable values, non-orthogonal rotations have been used (Imbrie and Van Andel, 1962; Miesch, 1976). The "oblique" rotation used by Imbrie and Van Andel (1962) involves the rotation of each varimax end-member to the simple "nearest" to it in composition (Figure 2c). Since each varimax axis is rotated to a sample contained in the data set, the end-member factors are obviously constrained to have realistic compositions. However, in order for this technique to adequately describe the data, some samples in the data set must be pure end-members. For fine-grained bottom sediment this is generally not the case. The use of an oblique projection can rotate orthogonal varimax factors until they coincide with the most divergent vectors. In this way, the others are defined as proportions of these objects. This is accomplished by constructing a $V_{(q \times q)}$ matrix that contains the highest absolute values of the varimax weights in each object column. The oblique projection matrix is given by

$$C = FV^{-1} \quad (5)$$

where V^{-1} is the inverse of V . The row vectors of the C matrix furnishes the proportional contributions of all the objects in terms of the reference objects. To recalculate these values in terms of the original data, it is necessary to divide each column vector of matrix C by the vector length of the corresponding object. This denormalizes the column vectors of the C matrix. If the number of analytes in the contaminant mixtures is unknown, oblique rotation permits its determination. The use of too many factors (or columns in V) results in physically meaningless concentration values in matrix C . In fact, many of these values are negative.

Factor analysis has not often been used to determine the actual composition of end-member sources in complex mixtures (Leinen and Piasias, 1984), because transformations of the original data variables during the statistical analysis result in negative factor scores for some variables and negative concentrations of some variables in the end-member. Thus, in the present paper the rotation technique (Figure 2d) proposed by Leinen and Piasias (1984) combined with an extended Q-mode factor analysis (Miesch, 1976) is used to determine chemically reasonable end-member compositions from the MM data set. This rotation scheme does not require the hypothesis of having sampled pure end-members (Full et al., 1981), but does assume true end-member compositions lie between the composition identified by Q-mode factor analysis (which forms a set of orthogonal axes) and the best known statistical parameter within a data set, i.e. the vector of mean composition (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002; Leinen and Piasias, 1984). Generally, the end-member compositions are found by rotating, one at a time, each varimax axis toward the mean vector until the composition of the rotated axis is chemically reasonable (i.e., all variable concentrations are greater than or equal to zero,

Figure 2d). These relations are accomplished by the following steps (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002):

- (1) Determine the varimax representation of the mean composition of the data. Let M denote the varimax vector of the mean composition.
- (2) Set a tolerance value defining "zero concentration" for each variable. This tolerance value is equal to the precision of the analysis or measurement of the variables. Any negative variable concentration within its tolerance value of zero will be considered equal to zero. For the application of this technique to the present Portland Harbor sediment data set, the tolerances were set equal to the precision with which the contaminants were determined.
- (3) For each varimax axis determine whether the variable concentrations are chemically acceptable (i.e., non-negative). If not, rotate the varimax vector toward the mean vector in steps scaled such that 100 steps are required to rotate the vector in question to the mean vector (Figure 2d). Algebraically this can be expressed as:

$$(1 - a)V_i + aM = E_i \quad (6)$$

where $a = 0.01, 0.02, \dots, 1.0$; V_i is the i -th varimax axis and E_i is the i -th end-member. After each step, the composition of E_i is tested to determine whether this new vector is within the acceptance region. After the rotation is complete, all negative concentrations smaller than the tolerance values are set to zero. The varimax representation and mean composition is determined from the total amount of information contained in the data set.

The criteria for choosing the number of end-members (Aboul-Kassim and Simoneit, 2001) used to model the Portland Harbor bottom sediment MM data were: (a) at least 90% of the variance in the data set must be explained by the sums of squares of the end-members, (b) all end-member factors that explained less than 2% of the total variance were rejected, and (c) all end-members which did not have a coherent distribution when mapped or plotted were rejected.

2.4.2. Linear Programming Technique

After identifying the end-member composition using this objective approach (Figure 2d), a linear programming technique (LPT) was used to determine the abundance of each end-member in each sediment sample (Aboul-Kassim, 1998; Aboul-Kassim and Williamson, 2002). This LPT utilized the inverse technique (Dymond, 1981) to providing a better fit of the observed multi-tracer environmental forensic MM data set with respect to the end-member compositions.

Once the number and composition of end-members were determined, the next step was to obtain a quantitative estimate of the relative amount of each end-member in each sediment sample. Because bottom sediments of the Portland Harbor are considered as mixtures of complex organic com-

pounds, the bulk composition of each sample from the harbor is assumed to consist of some linear combination of end-member compositions, each sediment sample was represented mathematically as a system of n equations (n = the number of individual contaminant variables used to identify the compositional end-members for sediment samples) in m unknowns (m = the number of major compositional end-members that are present) of the form:

$$\begin{aligned} S_{p1} &= K_1 E_{1p1} + K_2 E_{2p1} + \dots + K_m E_{mp1} + R_{p1} \\ S_{p2} &= K_1 E_{1p2} + K_2 E_{2p2} + \dots + K_m E_{mp2} + R_{p2} \\ &\vdots \\ S_{pn} &= K_1 E_{1pn} + K_2 E_{2pn} + \dots + K_m E_{mpn} + R_{pn} \end{aligned} \quad (7)$$

Where $S_{p1}, S_{p2}, \dots, S_{pn}$ = the measured concentrations of organic contaminant variables $p_{1,2} \dots p_n$ in the sediment sample; E_{1p1}, \dots, E_{mpn} = the concentrations of organic contaminant variables $p_1 \dots p_n$ in the compositional end-members $E_1 \dots E_m$ determined by factor analysis; $K_1, K_2 \dots K_n$ = unknowns whose magnitude for each sediment sample reflect the relative contributions of each compositional end-member in that sediment sample; and, $R_{p1}, R_{p2}, \dots, R_{pn}$ = residual terms reflecting the fact that each equation is in exact due to sampling and/or analytical error. These systems of equations are usually over-determined ($n > m$) in organic geochemical partitioning models, thus optimum solutions can be obtained using linear programming methods (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002). The major advantage of obtaining a linear programming solution is that certain physical constraints can be incorporated into the mathematical calculations. For example, the linear programming solution specifies that no compositional end-member can have a negative contribution to the total composition of sediment samples (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002).

The residual terms associated with each system of equations represent the difference between the linear programming estimate and the actual concentration of each organic contaminant in the bottom sediment sample. The optimum solution for each system of equations is that for which the residual terms are minimized. Since a perfect modeling solution would account for 100% of the measured concentration for each organic contaminant, the validity of the present environmental forensic MM model was evaluated by calculating a mean residual percent of each contaminant (i.e., the mean residual for each contaminant divided by the mean contaminant concentration) (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001). Thus, the use of linear programming technique partitioning helped correct the initial end-member compositions of bottom sediments and their abundance to better fit the observed multivariate data set, as well as to specify and select the compositions of the end-members.

3. Results and discussion

3.1. Molecular Markers

Molecular markers (MMs) are organic compounds, originating from various sources such as petroleum, coal, and wood via their combustion products. They are detected in the environment with structures suggesting an unambiguous link with known contemporary products. MMs are specific indicator compounds that can be utilized for genetic source correlation (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002, Dymond, 1981; Fraser and Lakshmanan, 2000; Hedges and Prahl, 1993; Kavouras et al., 1998; Kvenvolden et al., 1985; Kumata et al., 2002; Nolte et al., 1999; Philp, 1993; Standley et al., 2000; Zumberge, 1993). Such organic molecules are characterized by their restricted occurrence, source specificity, molecular stability, and suitable concentration for analytical detection (Aboul-Kassim and Simoneit, 2001). MMs have wide spread applications in organic compounds characterization and source identification (e.g., Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1996; 2001; Aboul-Kassim and Williamson, 2002, Curiale, 1993; Fraser and Lakshmanan, 2000; Hedges and Prahl, 1993; Kavouras et al., 1998; Kumata et al., 2002; Nolte et al., 1999; Philp, 1993; Porte et al., 2000; Salau et al., 1997; Standley et al., 2000; Zakaria et al., 2000; Zumberge, 1993). The following is an attempt to characterize the main hydrocarbon composition of bottom sediments sampled from the Portland Harbor.

3.1.1. Aliphatic Hydrocarbons

The aliphatic hydrocarbon MM suites examined for all bottom sediments showed the presence of several organic contaminants such as n -alkanes, isoprenoids, tri- and tetracyclic terpanes, hopanes, steranes, and diasteranes. The following sections discuss the various MMs present, both qualitatively and quantitatively. Table 1 shows the average MM composition of the Portland Harbor bottom sediments relative to both weight and TOC, respectively. Normal alkanes were present in all sediments, ranging in carbon chain length from C_{16} - C_{38} (Table 1). The average total n -alkane concentrations relative to dry weight and TOC representing the harbor environment amounted to 21.3 $\mu\text{g/g}$ and 3.3 mg/g , respectively. Figure 3a shows an example of GC-MS fingerprint of aliphatic hydrocarbon distributions representing the Portland Harbor bottom sediments.

The identification of the homologous n -alkanes in the aliphatic hydrocarbon fractions for different sediment samples allowed the determination of both carbon preference index (CPI) and C_{max} for each sample, which give supportive evidence for the relative incorporation of various aliphatic hydrocarbon sources. The CPI, a measure of biologically synthesized n -alkanes (Aboul-Kassim and Simoneit, 1995, 2001; Simoneit et al., 1990), indicates the relative contributions of n -alkanes from natural (CPI > 1) compared to anthropogenic (petroleum and industrial pollution; CPI \leq 1) sources. In the current EAIA project, the CPI calculated

according to Bray and Evans (1961) uses the same odd carbon number n -alkane concentrations in both ratios and the even carbon number concentrations in the denominator are shifted in one ratio versus the next (Table 2, ratio # 1). The average CPI value (i.e., 1.09) characteristic for the Portland Harbor sediments is shown in Table 2.

The determination of the C_{max} for every sediment sample also gave an indication of the relative source input, where a $C_{max} > C_{27}$ for n -alkanes reflects the incorporation of higher plant wax and C_{max} at lower carbon numbers indicates a major input from petrochemical sources (Aboul-Kassim, 1998; Aboul-Kassim and Simoneit, 1995, 1996, 2001; Simoneit, 1982). Generally, the dominant C_{max} determined for the n -alkanes of the various sediment samples are $< C_{27}$ with pristane predominance in many samples (Table 2), indicating a petrochemical source for most samples, with a minor terrestrial contribution.

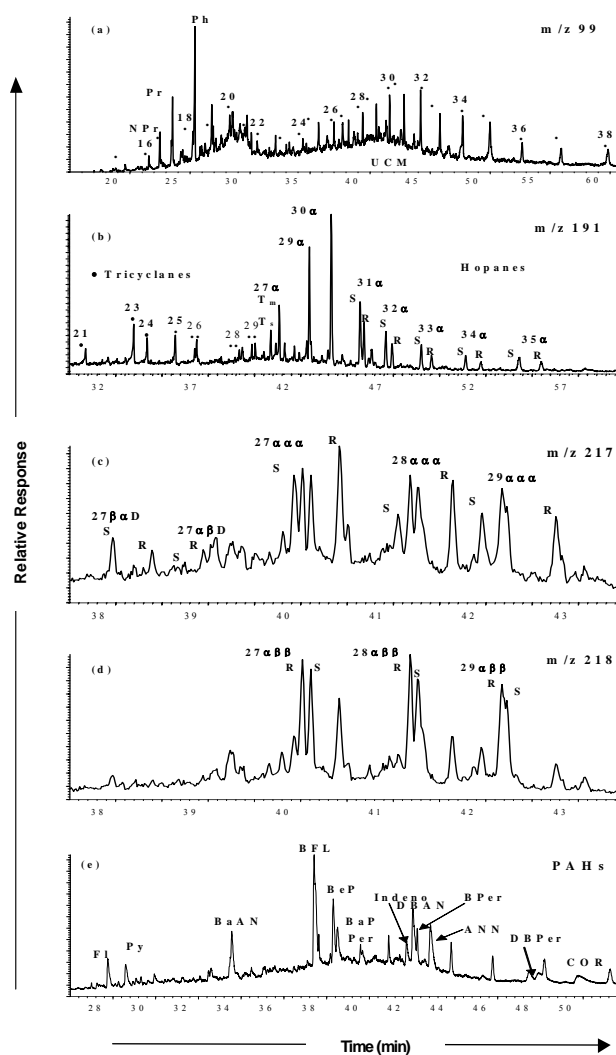
Isoprenoids such as pristane (2,6,10,14-tetramethylpentadecane, Figure 4, Structure II), phytane (2,6,10,14-tetramethylhexadecane, Figure 4, Structure III) and lower molecular weight homologs are geologic alteration products of phytol and are not primary constituents of most terrestrial biota (Aboul-Kassim and Simoneit, 1995, 1996, 2001; Peters and Moldowan, 1993; Rowland, 1990). Isoprenoid hydrocarbons (Table 1, Figure 3a) were present in most sediment samples mainly as pristane (average 88.2 $\mu\text{g/g}$ OC) and phytane (average 176.4 $\mu\text{g/g}$ OC), which confirms the petroleum related origin of the n -alkanes and UCM (e.g., Aboul-Kassim and Williamson, 2002; Peters and Moldowan, 1993; Simoneit, 1986, 1990; Simoneit and Kaplan, 1980). The distribution of these isoprenoids and their ratios (Table 2) for all samples points to a petrochemical input with a common source for bottom sediments.

Tricyclic terpanes (Figure 4, Structure IV), important geochemical tracers occurring in most Portland Harbor bottom sediments, range from $C_{19}H_{34}$ to $C_{45}H_{86}$ (Aboul-Kassim and Williamson, 2002; Aquino Neto et al., 1981, 1982; Peters and Moldowan, 1993; Moldowan and Seifert, 1983). The tricyclic terpane series (key ion m/z 191) is present in all sediment samples and ranges from $C_{19}H_{42}$ to $C_{29}H_{52}$, no C_{22} , with a C_{23} predominance (average 3.07 $\mu\text{g/g}$ OC, Table 1). The occurrence and variation in the relative distribution of the homologs of this series, as determined by GC-MS, in these sediment samples makes them useful tracers for petroleum source impact (Figure 3b).

Tetracyclic terpanes (Figure 3b), another group of MM detected in the harbor sediments, are derivatives of the hopanes (Aquino Neto et al., 1981). Both 17,21- and 8,14-*seco*-hopanes (Figure 4, Structures V and VI, respectively) are found in fossil fuels. The 17,21-*seco*-hopanes were proposed to be derived from either thermocatalytic degradation of hopane precursors, microbial ring opening of hopanoids during early diagenesis, or cyclization of squalene to ring-D (Trendel et al., 1982). The 8,14-*seco*-hopanes have ring-C opened hopane structures (Peters and Moldowan, 1993). The tetracyclic terpanes for the Portland Harbor sediments were comprised of a C_{24} -(17,21-

seco-hopane, i.e. E-norhopane) and C_{28} and C_{29} -(8,14-*seco*-hopanes) (Aboul-Kassim and Williamson, 2002).

Pentacyclic triterpane hydrocarbons (Figure 2b) from petroleum found in different samples from the Portland Harbor were the 17 α (H),21 β (H)-hopanes (Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002; Simoneit, 1986).



Notes: ¹Fl = fluoranthene, Py = pyrene, BaAN = benz[a]anthracene, BFL = benzo[b,k]fluoranthene, BeP = benzo[e]pyrene, BaP = benzo[a]pyrene, Per = perylene, Indeno = indeno[1,2,3-cd]pyrene, DBAN = dibenz[a,h]anthracene, BPer = benzo[ghi]perylene, ANN = anthanthrene, DBPer = dibenzo[ghi]perylene, Cor = coronene.
²(a) n -alkanes, m/z 99 (Pr = pristane, Ph = phytane, Npr = norpristane, UCM = unresolved complex mixture, numbers over peaks indicate carbon numbers); (b) hopane series, m/z 191; (c) $\alpha\alpha\alpha$ sterane series, m/z 217; (d) $\alpha\beta$ sterane series, m/z 218; and (e) PAH composition.

Figure 3. Typical GC-MS fingerprints representing the Portland Harbor contaminated sediments.

Table 1. Mean Aliphatic and Aromatic Hydrocarbon Compositions Relative to Dry Weight and TOC for the Portland Harbor Bottom Sediments

#	Compounds Name	Chemical Structures	MW	Mean Concentrations Relative to		ID*
				Dry Weight	TOC	
Aliphatic Hydrocarbons						
<i>n</i> -Alkanes (µg/g):						
1	<i>n</i> -hexadecane	C ₁₆ H ₃₄	226	0.0	0.7	a
2	<i>n</i> -heptadecane	C ₁₇ H ₃₆	240	0.2	33.8	a
3	<i>n</i> -octadecane	C ₁₈ H ₃₈	254	0.1	20.3	a
4	<i>n</i> -nonadecane	C ₁₉ H ₄₀	268	0.2	27.0	a
5	<i>n</i> -eicosane	C ₂₀ H ₄₂	282	0.3	47.3	a
6	<i>n</i> -heneicosane	C ₂₁ H ₄₄	296	0.6	87.8	a
7	<i>n</i> -docosane	C ₂₂ H ₄₆	310	0.7	101.3	a
8	<i>n</i> -tricosane	C ₂₃ H ₄₈	324	1.0	155.4	a
9	<i>n</i> -tetracosane	C ₂₄ H ₅₀	338	1.2	189.2	a
10	<i>n</i> -pentacosane	C ₂₅ H ₅₂	352	1.5	229.7	a
11	<i>n</i> -hexacosane	C ₂₆ H ₅₄	366	1.7	256.7	a
12	<i>n</i> -heptacosane	C ₂₇ H ₅₆	380	2.0	310.8	a
13	<i>n</i> -octacosane	C ₂₈ H ₅₈	394	2.2	337.8	a
14	<i>n</i> -nonacosane	C ₂₉ H ₆₀	408	2.3	358.1	a
15	<i>n</i> -triacontane	C ₃₀ H ₆₂	422	2.0	304.0	a
16	<i>n</i> -hentriacontane	C ₃₁ H ₆₄	436	2.0	304.0	a
17	<i>n</i> -dotriacontane	C ₃₂ H ₆₆	450	1.3	202.7	a
18	<i>n</i> -tritriacontane	C ₃₃ H ₆₈	464	0.9	135.1	a
19	<i>n</i> -tetracontane	C ₃₄ H ₇₀	478	0.4	67.6	a
20	<i>n</i> -pentatriacontane	C ₃₅ H ₇₂	492	0.2	33.8	a
21	<i>n</i> -hexatriacontane	C ₃₆ H ₇₄	506	0.2	27.0	a
22	<i>n</i> -heptatriacontane	C ₃₇ H ₇₆	520	0.1	20.3	a
23	<i>n</i> -octatriacontane	C ₃₈ H ₇₈	534	0.1	6.8	a
	Total <i>n</i> -alkanes (µg/g)			21.3	3257.1	
<u>Isoprenoids (µg/g):</u>						
24	2,6,10,14-tetramethylpentadecane (pristane)	C ₁₉ H ₄₀	254	13.5	88.2	a
25	2,6,10,14-tetramethylhexadecane (phytane)	C ₂₀ H ₄₂	268	27.0	176.4	a
	Total isoprenoids (µg/g)			40.5	264.6	
26	UCM (µg/g)			37.4	5723.4	
<u>Tricyclic Terpanes (ng/g):</u>						
27	C ₁₉ -tricyclic	C ₁₉ H ₃₄	262	10.4	67.7	b
28	C ₂₀ -tricyclic	C ₂₀ H ₃₆	276	31.3	204.6	b
29	C ₂₁ -tricyclic	C ₂₁ H ₃₈	290	10.4	67.7	b
30	C ₂₃ -tricyclic	C ₂₃ H ₄₂	318	115.2	752.2	b
31	C ₂₄ -tricyclic	C ₂₄ H ₄₄	332	78.7	513.7	b
32	C ₂₅ -tricyclic	C ₂₅ H ₄₆	346	68.1	444.5	b
33	C ₂₆ -tricyclic (S)	C ₂₆ H ₄₈	360	30.7	200.5	b
34	C ₂₆ -tricyclic (R)	C ₂₆ H ₄₈	360	31.6	206.0	b
35	C ₂₈ -tricyclic	C ₂₈ H ₅₀	388	36.8	239.9	c
36	C ₂₉ -tricyclic	C ₂₉ H ₅₂	402	57.5	375.4	c
	Total tricyclic terpanes (ng/g)			470.7	3072.3	
<u>Tetracyclic terpanes (ng/g):</u>						
37	C ₂₄ -tetracyclic (17,21-seco-hopane)	C ₂₄ H ₄₂	330	117.6	767.7	b
38	C ₂₈ -tetracyclic (18,14-seco-hopane)	C ₂₈ H ₅₀	386	39.8	259.7	b
39	C ₂₉ -tetracyclic (18,14-seco-hopane)	C ₂₉ H ₅₂	400	39.3	256.8	b
	Total tetracyclic terpanes (ng/g)			196.7	1284.2	
<u>Pentacyclic triterpanes (ng/g):</u>						
40	18α(H)-22,29,30-trisnorneohopane (Ts)	C ₂₇ H ₄₆	370	299.4	1954.5	b
41	17α(H)-22,29,30-trisnorhopane (Tm)	C ₂₇ H ₄₆	370	408.4	2665.8	b
42	17α(H),21β(H)-29-norhopane	C ₂₉ H ₅₀	398	1831.2	11952.9	b
43	17α(H),21β(H)-hopane	C ₃₀ H ₅₂	412	2301.0	15019.4	b
44	17α(H),21β(H)-homohopane (22S)	C ₃₁ H ₅₄	426	1041.6	6799.2	b
45	17α(H),21β(H)-homohopane (22R)	C ₃₁ H ₅₄	426	428.9	2799.8	b
46	17α(H),21β(H)-bishomohopane (22S)	C ₃₂ H ₅₆	440	742.0	4843.2	b
47	17α(H),21β(H)-bishomohopane (22R)	C ₃₂ H ₅₆	440	374.4	2444.2	b
48	17α(H),21β(H)-trishomohopane (22S)	C ₃₃ H ₅₈	454	599.1	3910.4	b
49	17α(H),21β(H)-trishomohopane (22R)	C ₃₃ H ₅₈	454	451.8	2949.2	b
50	17α(H),21β(H)-tetrakishomohopane (22S)	C ₃₄ H ₆₀	468	415.3	2710.9	b
51	17α(H),21β(H)-tetrakishomohopane (22R)	C ₃₄ H ₆₀	468	263.2	1718.2	b
52	17α(H),21β(H)-pentakishomohopane (22S)	C ₃₅ H ₆₂	482	435.6	2843.6	b
53	17α(H),21β(H)-pentakishomohopane (22R)	C ₃₅ H ₆₂	482	167.9	1095.9	b
	Total pentacyclic triterpanes (ng/g)			9759.9	63707.2	

<u>Diasteranes (ng/g):</u>						
54	13 α (H),17 β (H)-diacholestane (20S)	C ₂₇ H ₄₈	372	55.8	364.1	b,d
55	13 α (H),17 β (H)-diacholestane (20R)	C ₂₇ H ₄₈	372	59.9	390.9	b,d
	Total diasteranes (ng/g)			115.7	755.0	
<u>Steranes (ng/g):</u>						
56	5 α (H),14 α (H),17 α (H)-cholestane (20S)	C ₂₇ H ₄₈	372	22.5	147.0	b
57	5 α (H),14 β (H),17 β (H)-cholestane (20R)	C ₂₇ H ₄₈	372	53.7	350.7	b
58	5 α (H),14 β (H),17 β (H)-cholestane (20S)	C ₂₇ H ₄₈	372	32.9	214.8	b
59	5 α (H),14 α (H),17 α (H)-cholestane (20R)	C ₂₇ H ₄₈	372	22.6	147.3	b
60	5 α (H),14 α (H),17 α (H)-ergostane (20S)	C ₂₈ H ₅₀	386	22.5	147.0	b
61	5 α (H),14 β (H),17 β (H)-ergostane (20R)	C ₂₈ H ₅₀	386	52.0	339.4	b
62	5 α (H),14 β (H),17 β (H)-ergostane (20S)	C ₂₈ H ₅₀	386	53.7	350.7	b
63	5 α (H),14 α (H),17 α (H)-ergostane (20R)	C ₂₈ H ₅₀	386	19.1	124.4	b
64	5 α (H),14 α (H),17 α (H)-sitostane (20S)	C ₂₉ H ₅₂	400	43.3	282.7	b
65	5 α (H),14 β (H),17 β (H)-sitostane (20R)	C ₂₉ H ₅₂	400	79.7	520.4	b
66	5 α (H),14 β (H),17 β (H)-sitostane (20S)	C ₂₉ H ₅₂	400	64.1	418.4	b
67	5 α (H),14 α (H),17 α (H)-sitostane (20R)	C ₂₉ H ₅₂	400	28.0	182.6	b,d
	Total steranes (ng/g)			471.6	3078.4	
	Total Aliphatic Hydrocarbon (μg/g)			99.1	9245.3	
<u>Polycyclic Aromatic Hydrocarbons (PAHs)</u>						
<u>PAHs (ng/g):</u>						
68	Phenanthrene	C ₁₄ H ₁₀	178	27.9	181.8	a
69	Anthracene	C ₁₄ H ₁₀	178	0.0	0.0	a
70	Fluoranthene	C ₁₆ H ₁₀	202	210.5	1374.3	a
71	Pyrene	C ₁₆ H ₁₀	202	90.6	591.3	a
72	2,3-Benzofluorene	C ₁₇ H ₁₂	216	71.0	463.6	a
73	Benz(a)anthracene	C ₁₈ H ₁₂	228	420.6	2745.3	a
74	Chrysene/triphenylene	C ₁₈ H ₁₂	228	836.1	5457.8	a
75	Benzo(b+k)fluoranthene	C ₂₀ H ₁₂	252	967.1	6312.6	a
76	Benzo(e)pyrene	C ₂₀ H ₁₂	252	465.2	3036.9	a
77	Benzo(a)pyrene	C ₂₀ H ₁₂	252	196.2	1280.9	a
78	Perylene	C ₂₀ H ₁₂	252	54.0	352.2	a
79	Indeno(123-cd)pyrene	C ₂₂ H ₁₂	276	313.4	2045.9	b
80	Dibenz(ah)anthracene	C ₂₂ H ₁₄	278	109.4	714.2	a
81	Benzo(ghi)perylene	C ₂₂ H ₁₂	276	169.1	1103.8	a
82	Anthanthrene	C ₂₂ H ₁₂	276	0.0	0.0	a
83	Coronene	C ₂₄ H ₁₂	300	16.0	104.2	a
84	Dibenzo(ac)pyrene	C ₂₄ H ₁₄	302	165.6	1081.1	a
	Total PAHs (μ g/g)			4.1	26.9	
<u>Alkyl-substituted PAHs (ng/g):</u>						
85	3-Methylphenanthrene (3MP)	C ₁₅ H ₁₂	192	18.3	119.6	a
86	2-Methylphenanthrene (2MP)	C ₁₅ H ₁₂	192	16.8	109.7	a
87	9-Methylphenanthrene (9MP)	C ₁₅ H ₁₂	192	23.8	155.6	a
88	1-Methylphenanthrene (1MP)	C ₁₅ H ₁₂	192	26.6	173.6	a
89	Dimethylphenanthrene	C ₁₆ H ₁₄	206	228.4	1491.0	b
90	Trimethylphenanthrene	C ₁₇ H ₁₆	220	244.4	1595.4	b
91	Tetramethylphenanthrene	C ₁₈ H ₁₈	234	105.1	686.3	b
	Total Phenanthrene Series (ng/g)			663.6	4331.3	
92	Methylpyrene/fluoranthene	C ₁₇ H ₁₂	216			b
93	Dimethylpyrene/fluoranthene	C ₁₈ H ₁₄	230	303.4	1980.3	b
94	Trimethylpyrene/fluoranthene	C ₂₀ H ₁₆	244			d
	Total alkyl-202 (pyrene+fluoranthene) Series (ng/g)			303.4	1980.3	
95	Methyl-228	C ₁₉ H ₁₄	242	393.0	2565.1	b
96	Dimethyl-288	C ₂₀ H ₁₆	256	351.1	2291.6	b
97	Trimethyl-228	C ₂₁ H ₁₈	270	268.8	1754.3	b
	Total Alkyl-228 (BaAn+Chr+Triph) Series (ng/g)			1012.8	6611.0	
98	Methyl-252	C ₂₁ H ₁₄	266	522.7	3412.0	b
99	Dimethyl-252	C ₂₂ H ₁₆	280	598.2	3905.0	b
100	Trimethyl-252	C ₂₃ H ₁₈	294	433.4	2828.8	b
101	Tetramethyl-252	C ₂₄ H ₂₀	308	406.3	2651.9	b
	Total Alkyl-252 [B(b+k)fl+BeP+BaP+per] Series (ng/g)			1960.6	12797.7	
	Total Alkyl-substituted PAHs (μ g/g)			3.9	25.7	
	Total PAHs + alkyl PAHs (μg/g)			8.1	52.6	

*Aliphatic and polycyclic aromatic hydrocarbons molecular marker ratios for the Portland Harbor bottom sediments

Table 2. Aliphatic and Polycyclic Aromatic Hydrocarbons Molecular Marker Ratios for the Portland Harbor Bottom Sediments

#	Molecular Marker Ratios	Mean Values
Aliphatic Hydrocarbons Indices		
1	$CPI = \frac{1}{2} \left(\frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{24} + C_{26} + C_{28} + C_{30} + C_{32}} + \frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{26} + C_{28} + C_{30} + C_{32} + C_{34}} \right)$	1.092
2	$\left(\frac{U}{R} \right) = \left(\frac{\text{Concentration of unresolved complex mixture (UCM)}}{\text{Concentrations of } \sum \text{ resolved hydrocarbon peaks (mostly n-alkanes)}} \right)$	1.757
3	$\left(\frac{Pr}{Ph} \right) = \left(\frac{2,6,10,14\text{-tetramethylpentadecane (pristane)}}{2,6,10,14\text{-tetramethylhexadecane (phytane)}} \right)$	0.500
4	$\left(\frac{Pr}{C_{17}} \right) = \left(\frac{\text{Pristane}}{C_{17}} \right)$	0.400
5	$\left(\frac{Ph}{C_{18}} \right) = \left(\frac{\text{Phytane}}{C_{18}} \right)$	1.333
6	$C_{max} = \text{The highest n-alkane peak}$	27
7	$\left(\frac{Ts}{Tm} \right) = \left(\frac{18\alpha(H)\text{-}22,29,30\text{-trisorneohopane}}{17\alpha(H)\text{-}22,29,30\text{-trisnorhopane}} \right)$	0.733
8	$\text{Triplet Ratio} = \left[\frac{(C_{26}\text{-Tricyclic S}) + (C_{26}\text{-Tricyclic R})}{(C_{24}\text{-Tetracyclic})} \right]$	0.530
9	$\left(\frac{C_{23}Tri}{C_{30}\beta} \right) = \left(\frac{C_{23}\text{-tricyclic}}{17\alpha(H),21\beta(H)\text{-hopane}} \right)$	0.050
10	$\left(\frac{C_{29}\alpha}{C_{30}\beta} \right) = \left(\frac{17\alpha(H),21\beta(H)\text{-norhopane}}{17\alpha(H),21\beta(H)\text{-hopane}} \right)$	0.796
Homohopane Index (HHI)		
HHI is the ratio between the S and R epimer at C-22 for the 17 α (H)-homohopane series		
$(C_{31} - C_{35}) = \left[\frac{22S}{(22S+22R)} \right], \text{ as follows:}$		
11		C ₃₁ 0.708
12		C ₃₂ 0.665
13		C ₃₃ 0.570
14		C ₃₄ 0.612
15		C ₃₅ 0.722
Sterane Epimerization Indices		
Sterane epimerization parameter at C-20 is calculated for C ₂₉ for 5 α (H),14 α (H),17 α (H)-C ₂₉ -sterane and 5 α (H),14 β (H),17 β (H)-C ₂₉ -sterane as:		
16	$\alpha\alpha\alpha C_{29} \left[\frac{20S}{(20S+20R)} \right]$	0.608
17	$\alpha\beta\beta C_{29} \left[\frac{20S}{(20S+20R)} \right]$	0.696
PAH Indices		
18	$\left(\frac{\% \text{ Alkyl PAHs}}{\sum \text{ PAHs}} \right) = \left(\frac{\sum \text{ concentrations of all alkyl PAHs}}{\sum \text{ concentrations of all PAHs}} \right) \times 100$	48.93
19	$\left(\frac{P}{P+An} \right) = \left[\frac{\text{Phenanthrene}}{\text{anthracene} + \text{phenanthrene}} \right]$	1.000
20	$\left(\frac{B(a)An}{B(a)An+Chr+Triph} \right) = \left[\frac{\text{benz(a)anthracene}}{\text{benz(a)anthracene} + \text{chrysene} + \text{triphenylene}} \right]$	0.116
21	$\left(\frac{B(e)P}{B(e)P+B(a)P} \right) = \left[\frac{\text{benzo(e)pyrene}}{\text{benzo(e)pyrene} + \text{benzo(a)pyrene}} \right]$	0.703
22	$\left(\frac{Inpy}{Inpy+Bper} \right) = \left[\frac{\text{indeno(123-cd)pyrene}}{\text{indeno(123-cd)pyrene} + \text{benzo(ghi)perylene}} \right]$	0.650

The identification of these compounds is based primarily on their mass spectra and GC retention time in the key ion fragmentogram (Aboul-Kassim and Simoneit, 2001; Philp, 1985). The predominant analog in these samples (Figure 2b) is 17 α (H),21 β (H)-hopane (Figure 3, Structure VIII), with subordinate amounts of 18 α (H)-22,29,30-trisnorhopane (Ts), 17 α (H)-22,29,30 tris- norhopane (Tm), 17 α (H), 21 β (H)-29-norhopane, and minor concentrations of the 17 β (H), 21 α (H)-hopanes and the extended 17 α (H), 21 β (H)-hopanes ($> C_{31}$) (Aboul-Kassim and Williamson, 2002). The distributions of the 17 α (H)-hopane series (Figure 2b) are found for petroleum and petrochemical industry, confirming that major input for the different harbor sediments (Aboul-Kassim and Simoneit, 2001; Aboul-Kassim and Williamson, 2002). In typical petroleum and their products, the extended 17 α (H), 21 β (H)-hopane homologs $> C_{31}$ have the epimers at C-22 at an equilibrium ratio $[S/(S+R)]$ of 0.6 (homohopane index) (Aboul-Kassim and Simoneit, 2001; Seifert and Moldowan, 1978). The homohopane index (Table 2, ratios 11-15) for these bottom sediments varies from 0.57 to 0.72.

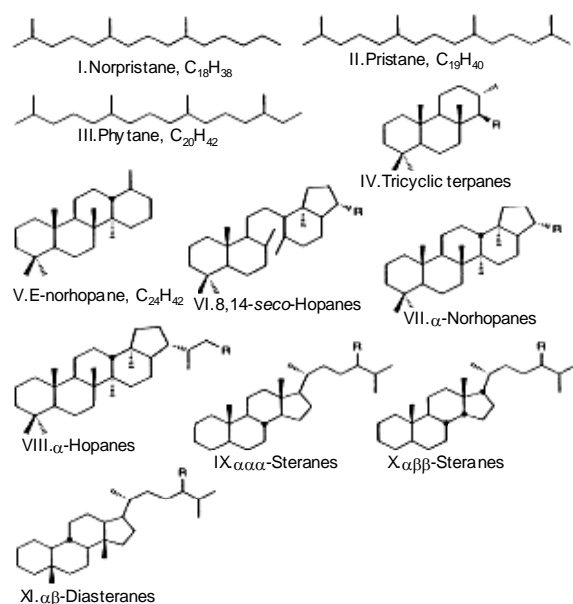


Figure 4. Chemical structures of some aliphatic hydrocarbons.

Steranes and diasteranes present in fossil fuels are useful molecular marker (MM) indicators for petroleum impact. These MMs were detected (e.g., Figures 3c,d) and quantified (Table 1) in the harbor samples relative to weight and carbon content. The steranes (Table 1) have mainly the 5 α (H),14 β (H),17 β (H)-configuration (Figure 4, Structure X), and a minor amount of the 5 α (H), 14 α (H),17 α (H)-configuration (Figure 4, Structure IX), with traces of 13 α (H),17 β (H)-diasteranes (Figure 4, Structure XI). The epimerization ratio at C-20 of the C₂₉ sterane (Table 2, ratios 16-17) is characteristic petroleum residues

for such bottom sediments (Aboul-Kassim and Simoneit, 2001).

Besides the chromatographically resolved compounds, another diagnostic (i.e., genetic source) parameter representing the Portland Harbor aquatic sediments was the presence of unresolved complex mixture (UCM) of branched and cyclic hydrocarbons (regional averages of 37.4 $\mu\text{g/g}$ dry weight and 5.7 mg/g OC, respectively). These UCMs were eluting between $n\text{-C}_{16}$ and $n\text{-C}_{33}$ in most bottom sediment (e.g., Figure 3a), but with different strengths among samples. The presence of UCM is interpreted here to derive from utilization of petroleum products (e.g., Aboul-Kassim and Simoneit, 2001; Dymond, 1981; Kennicutt II et al., 1994; Mazurek and Simoneit, 1983; Rogge et al., 1993a, b; Simoneit, 1985). The ratio between unresolved to resolved (U/R) hydrocarbons (average 1.76, Table 2) was also used as a criterion to assess anthropogenic input. A U/R value >1 reflects significant contamination by petroleum products (e.g., Aboul-Kassim and Simoneit, 1996, 2001; Aboul-Kassim and Williamson, 2002; Mazurek and Simoneit, 1983).

3.1.2. Polycyclic Aromatic Hydrocarbon (PAHs)

The second main class of compounds studied in the EAIA multi-tracer model consists of PAHs (Table 1). They are considered to be a class of chemical carcinogens and mutagenic pollutants (Figure 5, Structures I-XIX), derived from anthropogenic sources, such as vehicular exhaust, emissions from coke ovens, refining, and other processes involving high temperature pyrolytic reactions and incineration (Aboul-Kassim and Simoneit, 2001; Simoneit, 1984, 1998; Simoneit and Aboul-Kassim, 1995).

The mean concentrations of individual PAHs in different sediment samples in terms of weight and TOC are given in Table 1, and account for all detectable PAH by GC and GC-MS. Regardless of sediment samples, the PAH mixtures were similar with some alkylated homologs observed mainly for phenanthrene, pyrene, perylene and benz(a)anthracene with concentration differences among samples. The PAHs showed a predominance of fluoranthene, benz(a)anthracene, chrysene/triphenylene, benzo(b+k) fluoranthene, benz(e)-pyrene and benz(a)pyrene for bottom sediments (Table 1). Figure 3e shows an example of PAH GC-MS tracer representing the various sediment samples in the Portland Harbor.

Generally the proportion of alkylated to parent PAH depends on the combustion temperature (Simoneit, 1998). Thus, coal and wood smokes contain a phenanthrene mixture maximizing at the parent PAH with an exponential drop to the C₄-homologs. In contrast, vehicular emissions exhibit a pattern of low amounts of phenanthrene and maximum at the C₁-homologs, and petroleum input is characterized by a distribution increasing uniformly from less to more alkylated homologs up to C₅ and greater. For most bottom sediments, the alkylated PAHs of the phenanthrene, fluoranthene/pyrene, -228 and -252 series maximize at C₂, C₂, C₁₋₂ and C₂, respectively (Table 1). Relatively high concentrations of 2- and 3-methylphenanthrene (MP) compared to 1-

and 9-MP were observed for most bottom sediments (Table 1) indicating thermogenic alteration. This can be explained in terms of the rearrangement of the MP, favoring the thermodynamically more stable 2- and 3-positions at high temperatures (Aboul-Kassim and Simoneit, 2001; Simoneit, 1998).

Thus, coupling CPI, C_{max} , UCM, quantitation of MMs, different organic geochemical parameters and PAHs allowed the definition of the main sources of the MMs (petrogenic vs. thermogenic/pyrogenic) characteristic for Portland Harbor sediments. However, these analyses presented only the assessments of the different sources of MMs in the bottom sediments and not their source strengths (cf. the statistical/mathematical modeling part).

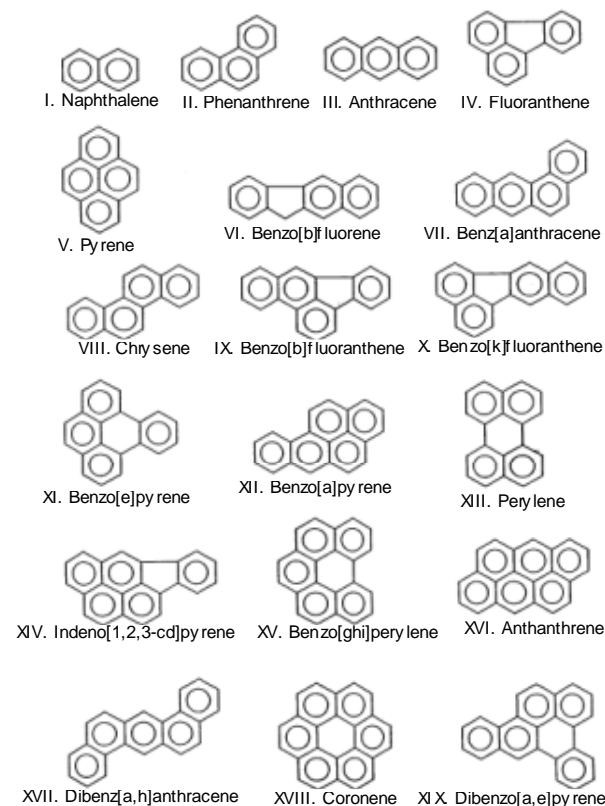


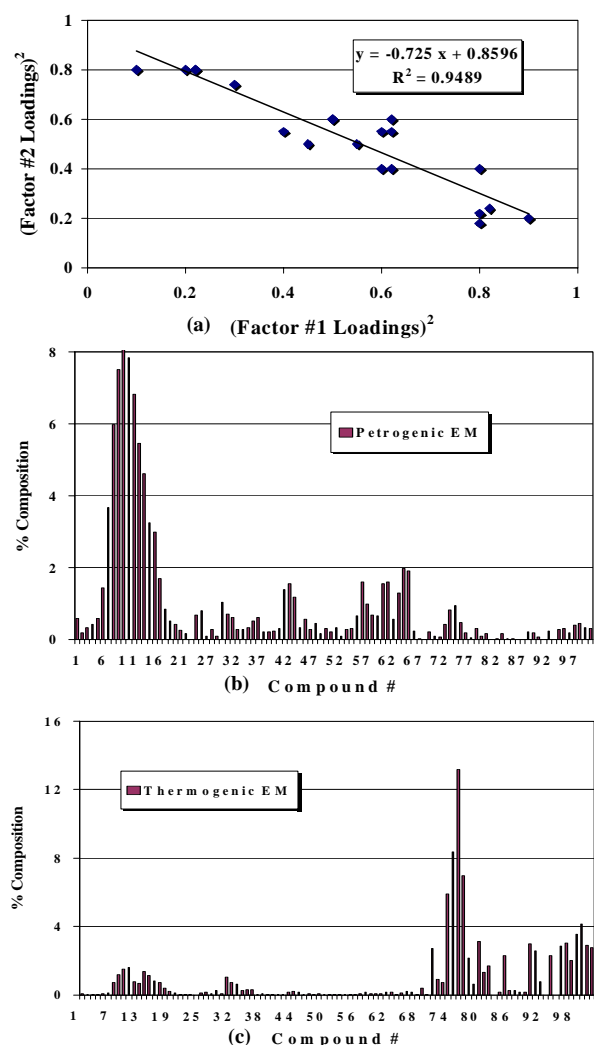
Figure 5. Chemical structures of some polycyclic aromatic hydrocarbons

3.2. Data Interpretation and Source Confirmation

The EAIA environmental forensic multi-tracer model for the various sediment samples generated a large amount of data, necessitating the use of statistical/mathematical modeling techniques to group the data into significant groups and reduce them into a number of factors, which represent in an organic geochemical partitioning sense, the combined effects of several chemical processes or factors.

In the case of Q-mode factor analysis of the EAIA data set, the first result yields two significant principal factors loading scores, providing information about sample varia-

tion of about 72.3% and 24.3%, respectively. When individual MM compounds of this factor-loading matrix are squared, the sum of the squared loadings for all factors of a particular sediment sample equals 1.00 (i.e. communality; which is the proportion of the total variance in particular sediment sample that is explained by those factors). The individual squared loading of one factor represents the fraction of the sediment sample, which that factor contributes to the aquatic sediment sample (e.g., if a sample has a factor 1 loading of 0.4 then $(0.4)^2 = 0.16$ or 16% of the sediment sample contamination is from factor 1) (Aboul-Kassim and Simoneit, 2001). Thus, the distribution of the various factors in each sediment sample was obtained. A plot is used to observe associations (Aboul-Kassim and Simoneit, 2001; SAS, 1997; Rapp, 1991) between samples (groupings; Figure 6a).



Notes: (a) factor loading squared; (b) first “petrogenic” end-member and (c) second “thermogenic/pyrogenic” end-member.

Figure 6. Genetic source partitioning model after extended Q-mode factor analysis and linear programming technique representing the Portland Harbor contaminated sediments.

Most sediment samples plot near the binary mixing line (a line from factor 1 = 1.00, factor 2 = 0 to factor 1 = 0, factor 2 = 1.00), indicating that two main factors can explain the majority of the composition of the samples (Aboul-Kassim and Simoneit, 2001). The second result is from squaring individual elements of the factor score matrix yielding the sum for a particular factor equal to 1.00. The proportion, which an individual MM class contributes to the total composition of an end-member, is determined by dividing the absolute values of all MMs scores for that factor by the sum of the absolute values of all the scores for that factor (Aboul-Kassim and Simoneit, 2001).

After using the rotation proposed by Leinen and Piasis (1984) and linear programming techniques, the two end-member (EM I & EM II) compositions were obtained (see Figure 6b,c). In order to assign the origin (source) and interpret the observed factor data, both statistical/mathematical modeling and MM approaches were applied. End-member #1 is dominated mainly by *n*-alkanes, regular isoprenoid hydrocarbons, tri- and tetracyclic series, the hopane series with a predominance of 17 α (H), 21 β (H)-hopane and both C-22 S/R configurations for the homologs > C₃₁, and the diasterane/sterane series. In addition, a minor contribution of different unsubstituted and substituted PAHs recorded 4.09% and 3.03%, respectively (Figure 6b). End-member #2 is dominated mainly by a group of unsubstituted and substituted PAHs with a minor contribution of terrestrial aliphatic hydrocarbons. Fluoranthene, benz(a)anthracene, chrysene/triphenylene, benz(b+k)fluoranthene, benzo(e)pyrene, benzo(a)pyrene, indeno(123-cd)pyrene, and benzo(ghi)perylene were dominating the unsubstituted PAHs in the second end-member (see Figure 6c).

Based on the data analyses findings and the molecular markers (MMs) specific for bottom sediments collected from the Portland Harbor, two main sources can be assigned: (a) *Petrogenic Source*, represented by end-member 1, and (b) *Thermogenic/pyrogenic Source*, represented by end-member 2. Source assignment modeling for the vast MMs specific for bottom sediments is important in the present research. Its importance is because the environmental toxicity and chemodynamics of such contaminated sediment particles can be easily predicted by studying only such "statistically and experimentally" assigned fractions as well as their molecular compositions.

4. Conclusions

Bottom sediments collected from the Portland Harbor of Oregon (USA) contain various anthropogenic hydrocarbon compounds, which are specific to their emission sources. This anthropogenic components comprise mainly petroleum residues, confirmed by UCM, U/R, pristane, phytane, CPI, C_{max}, tricyclic (C₁₉ - C₂₉) and tetracyclic (C₂₄, C₂₈ and C₂₉ *seco*-hopanes) terpanes, triterpanes ($\alpha\beta$ -hopanes), steranes ($\alpha\beta\beta$ configuration with a minor amount of the $\alpha\alpha\alpha$), diasteranes, and thermogenic/pyrogenic PAHs, ranging from

fluorene to coronene with minor alkyl-PAH series.

The multivariate statistical analyses, including both extended Q-mode factor analysis and linear programming technique, reduced the hydrocarbon data set into two significant end-members (sources), explaining 96.6% of the variation among the sediment samples. These multivariate techniques represented a useful method for end-member source confirmation, representing petrogenic (72.3%) and thermogenic/pyrolytic (24.3%) sources. The findings indicate that this approach is a cost-effective approach for environmental analysis and impact assessment.

Acknowledgements. The authors acknowledge the financial support (Circular #U0226A) from the Superfund Basic Research Center (SBRC) of Oregon Health Sciences University (OHSU) through the National Institute of Environmental Health Sciences (NIEHS).

References

- Aboul-Kassim, T.A.T. (1998). *Characterization, Chemodynamics and Environmental Impact Assessment of Organic Leachates from Complex Mixtures*, Ph.D. Dissertation, College of Engineering, Department of Civil, Construction and Environmental Engineering, Oregon State University, Corvallis, OR, USA.
- Aboul-Kassim, T.A.T. and Simoneit, B.R.T. (1995). Petroleum hydrocarbon fingerprinting and sediment transport assessed by molecular biomarker and multivariate statistical analyses in the Eastern Harbor of Alexandria. *Mar. Poll. Bull.*, 30, 63-73.
- Aboul-Kassim, T.A.T. and Simoneit, B.R.T. (1996). Lipid geochemistry of surficial sediments from the coastal environment of Egypt. I-Aliphatic hydrocarbons-characterization and sources. *Mar. Chem.*, 54, 135-158.
- Aboul-Kassim, T.A.T. and Simoneit, B.R.T. (2001). Pollutant-solid phase interactions: Mechanisms, chemistry and modeling, in *The Handbook of Environmental Chemistry, Water Pollution Series*, Springer-Verlag, Heidelberg, Germany, Vol. 5/Part E, pp. 430.
- Aboul-Kassim, T.A.T. and Williamson, K.J. (2002). *Environmental Analysis and Impact Assessment of Portland Harbor Contaminated Sediments*, Final Progress Report: Summary of Approaches, Methodology, Laboratory Results and Model Development, Circular #U0226A, the Superfund Basic Research Center (SBRC), Oregon Health and Science University (OHSU) and the National Institute of Environmental Health Sciences (NIEHS), Department of Civil, Construction and Environmental Engineering, Oregon State University, OR, USA, pp. 191, April 2002.
- Aquino Neto, F.R., Restle, A., Connan, J., Albrecht, P. and Ourisson, G. (1982). Novel tricyclic terpanes (C₁₉-C₂₀) in sediments and petroleum. *Tetrahedron Lett.*, 23, 2027-2030.
- Aquino Neto, F.R., Trendel, J.M., Restle, A., Connan, J. and Albrecht, P. (1981). Occurrence and formation of tricyclic and tetracyclic terpanes in sediments and petroleum, in Bjorøy et al. (Eds.), *Advances in Organic Geochemistry*, John Wiley & Sons, Chichester, UK, pp. 695-676.
- Bray, E.E. and Evans, E.D. (1961). Distribution of *n*-paraffins as a clue to recognition of source beds. *Geochim. Cosmochim. Acta.*, 22, 2-15.
- Curiale, J.A. (1993). Oil to source rock correlation: concepts and case studies, in M.H. Engel and S.A. Macko (Eds.), *Organic Geochemistry-Principles and Applications*, Plenum Press, NY, USA, pp. 473-490.

- Dymond, J. (1981). The geochemistry of Nazca Plate surface sediments: An evaluation of hydrothermal, biogenic, detrital and hydrogenous sources, in Kulm et al. (Eds.), *Nazca Plate - Crustal Formation and Andean Convergence*, Geol. Soc. Amer. Mem., 154, pp. 133-174.
- Eglinton, G. and Calvin, M. (1967). Chemical fossils. *Sci. Am.*, 216, 32-43.
- Fraser, M.P. and Lakshmanan, K. (2000). Using levoglucosan as a molecular marker for the long-range transport of biomass combustion aerosols. *Environ. Sci. Technol.*, 34(21), 4560-4564.
- Full, W.E., Ehrlich, R. and Klován, J.E. (1981). Extended Q-mode objective definition of external end-members in the analysis of mixtures. *J. Math. Geol.*, 13, 331-334.
- Grant, A. (1990). Multivariate statistical analyses of sediment geochemistry. *Mar. Poll. Bull.*, 21, 297-299.
- Heath, G.R. and Dymond, J. (1981). Metalliferous sediment deposition in time and space: East Pacific Rise and Bauer Basin, northern Nazca Plate, in Kulm et al. (Eds.), *Nazca Plate - Crustal Formation and Andean Convergence*, Geol. Soc. Amer. Mem., 154, pp. 175-198.
- Hedges, J.I. and Prahl, F.G. (1993). Early diagenesis: consequences for applications of molecular markers, in M.H. Engel and S.A. Macko (Eds.), *Organic Geochemistry - Principles and Applications*, Plenum Press, NY, USA, pp. 237-253.
- Imbrie, J. and Purdy, E.G. (1962). Classification of modern Bahamian carbonate sediments, in W.E. Ham (Ed.), *Amer. Assoc. Petrol. Geol.*, Memoir 1, Tulsa, OK, USA, pp. 253-272.
- Imbrie, J. and Van Andel, T.H. (1962). Vector analysis of heavy mineral data. *Geol. Soc. Amer. Bull.*, 75, 1131-1156.
- Kavouras, I.G., Stratigakis, N., Stephanou, E.G. (1998). Iso- and Anteiso-Alkanes: specific tracers of environmental tobacco smoke in indoor and outdoor particle-size distributed urban aerosols. *Environ. Sci. Technol.*, 32(10), 1369-1377.
- Kennicutt II, M.C., Wade, T.L., Presley, B.J., Requejo, A.G., Brooks, J.M. and Denoux, G. (1994). Sediment contaminants in Casco Bay, Maine: Inventories, sources, and potential for biological impact. *Environ. Sci. Technol.*, 28, 1-15.
- Klován, J.E. and Imbrie, J. (1971). An algorithm and Fortran-IV program for large scale Q-mode factor analysis and calculation of factor scores. *Math. Geology*, 3, 61-76.
- Kumata, H., Yamada, J., Masuda, K., Takada, H., Sato, Y., Sakurai, T. and Fujiwara, K. (2002). Benzothiazolamines as tire-derived molecular markers: Sorptive behavior in street runoff and application to source apportioning. *Environ. Sci. Technol.*, 36(4), 702-708.
- Kvenvolden, K.A., Rapp, J.B. and Bourell, J.H. (1985). Comparison of molecular markers in crude oils and rocks from the north slope of Alaska, in L.B. Magoon and G.E. Claypool (Eds.), *Alaska North Slope Oil/Rock Correlation Study*, American Association of Petroleum Geologists, Tulsa, OK, USA, No. 20, pp. 593-617.
- Leinen, M. and Pisas, N. (1984). An objective technique for determining end-member compositions and for partitioning sediments according to their sources. *Geochim. Cosmochim. Acta.*, 48, 47-62.
- Malinowski, E.R. (1991). *Factor Analysis in Chemistry*, John Wiley & Sons, New York, NY, USA, pp. 326.
- Mazurek, M. and Simoneit, B.R.T. (1983). Characterization of biogenic and petroleum derived organic matter in aerosols over remote, rural and urban areas, in L.H. Keith (Ed.), *Identification and Analysis of Organic Pollutants*, Ann Arbor Science/Butterworth Publishers: Woburn, MA, USA, pp. 353-370.
- Miesch, A.T. (1976). Q-mode factor analysis of geochemical and petrologic data matrices with constant row sums. *U.S. Geol. Surv. Prof. Pap.*, 574G.
- Moldowan, J.M., Seifert, W.K. and Gallegos, E.J. (1983). Identification of an extended series of tricyclic terpanes in petroleum. *Geochim. Cosmochim. Acta.*, 47, 1531-1534.
- Nolte, C.G., Schauer, J.J., Cass, G.R. and Simoneit, B.R.T. (1999). Highly polar organic compounds present in meat smoke. *Environ. Sci. Technol.*, 33(19), 3313-3316.
- Peters, K.E. and Moldowan, J.M. (1993). *The Biomarker Guide, Interpreting Molecular Fossils in Petroleum and Ancient Sediments*, Prentice Hall, Englewood Cliffs, NJ, USA, pp. 363.
- Philp, R.P. (1985). *Fossil Fuel Biomarkers Methods in Geochemistry and Geophysics*, Elsevier, New York, NY, USA, vol. 23, pp. 292.
- Philp, R.P. (1993). Oil-oil and oil-source rock correlations: Techniques, in M.H. Engel and S.A. Macko (Eds.), *Organic Geochemistry - Principles and Applications*, Plenum Press, NY, USA, pp. 445-460.
- Porte, C., Biosca, X., Pastor, D., Sole, M. and Albaiges, J. (2000). The Aegean sea oil spill. 2. Temporal study of the hydrocarbons accumulation in bivalves. *Environ. Sci. Technol.*, 34(24), 5067-5075.
- Rapp, J.B. (1991). A statistical approach to the interpretation of aliphatic hydrocarbon distributions in marine sediments. *Chem. Geol.*, 93, 163-177.
- Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R. and Simoneit, B.R.T. (1993a). Sources of fine organic aerosol. 2. Noncatalyst and catalyst-equipped automobiles and heavy-duty diesel trucks. *Environ. Sci. Technol.*, 27, 636-651.
- Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R. and Simoneit, B.R.T. (1993b). Characterization of biogenic and petroleum derived organic matter. *Environ. Sci. Technol.*, 27, 2700-2711.
- Rowland, S.J. (1990). Production of acyclic isoprenoid hydrocarbons by laboratory maturation of methanogenic bacteria. *Org. Geochem.*, 15, 9-16.
- Salau, J.S.I., Tauler, R., Bayona, J.M. and Tolosa, I. (1997). Input characterization of sedimentary organic contaminants and molecular markers in the northwestern Mediterranean sea by exploratory data analysis. *Environ. Sci. Technol.*, 31(12), 3482-3490.
- SAS (1997). *SAS/STAT™ User's Guide*, Release 9.0 Edition, SAS Institute Inc., Cary, NC, USA, pp. 1028.
- Seifert, W.K. and Moldowan, J.M. (1987). Applications of steranes, terpanes, and monoaromatics to the maturation, migration, and source of crude oils. *Geochim. Cosmochim. Acta.*, 42, 77-95.
- Simoneit, B.R.T. (1982). Composition, sources and transport of organic matter to marine sediments-The organic geochemical approach, in J.A.J. Thompson and W.D. Jamieson (Eds.), *Marine Chemistry into the Eighties*, National Research Council of Canada, Ottawa, Canada, pp. 82-112.
- Simoneit, B.R.T. (1984). Organic matter of the troposphere-III characterization and sources of petroleum and pyrogenic residues in aerosols over the western United States. *Atmos. Environ.*, 18, 51-67.
- Simoneit, B.R.T. (1985). Application of molecular marker analysis to vehicular exhaust for source reconciliations. *Int. J. Environ. Analyt. Chem.*, 22, 203-233.
- Simoneit, B.R.T. (1986). Cyclic terpenoids of the geosphere, in R.B. Johns (Ed.), *Biological Markers in the Sedimentary Record*, Elsevier, NY, USA, pp. 43-99.
- Simoneit, B.R.T. (1998). Biomarker PAHs in the environment, in A.H. Neilson (Ed.), *The Handbook of Environmental chemistry*, Springer-Verlag, Heidelberg, Germany, Chapt. 5, pp. 176-221.
- Simoneit, B.R.T. and Aboul-Kassim, T.A.T. (1995). *Detection of*

- Fuels on Water in a Port of an Estuary*, Final Progress Report, Pacific States Marine Fisheries Commission, Depoe Bay, OR, USA, pp. 112.
- Simoneit, B.R.T. and Kaplan, I.R. (1980). Triterpenoids as molecular indicators of paleoseepage in recent sediments of the southern California Bight. *Mar. Environ. Res.*, 3, 113-128.
- Simoneit, B.R.T., Cardoso, J.N. and Robinson, N. (1990). An assessment of the origin and composition of higher molecular weight organic matter in aerosols over Amazonia. *Chemosphere*, 21, 1285-1301.
- Standley, L.J., Kaplan, L.A. and Smith, D. (2000). Molecular tracers of organic matter sources to surface water resources. *Environ. Sci. Technol.*, 34(15), 3124-3130.
- Trendel, J.M., Restle, A., Connan, J. and Albrecht, P. (1982). Identification of a novel series of tetracyclic terpene hydrocarbons (C₂₄-C₂₇) in sediments and petroleum. *J. Chem. Soc.*, 304-325.
- Tysklind, M., Lundergen, K., Rappe, C., Eriksson, L., Jonsson, J., Sjöström, M. and Ahlberg, U.G. (1992). Multivariate characterization and modeling of polychlorinated dibenzo-p-dioxins and dibenzofurans. *Environ. Sci. Technol.*, 26, 1023-1030.
- Zakaria, M.P., Horinouchi, A., Tsutsumi, S., Takada, H., Tanabe, S. and Ismail, A. (2000). Oil pollution in the straits of Malacca, Malaysia: application of molecular markers for source identification. *Environ. Sci. Technol.*, 34, 1189-1196.
- Zumberge, J.E. (1993). Organic geochemistry of Estancia Vieja oils, Rio Negro Norte Block: Correlation with other Neuquen Basin, Argentina oils, in M.H. Engel and S.A. Macko (Eds.), *Organic Geochemistry - Principles and Applications*, Plenum Press, NY, USA, pp. 461-471.